

Fielded sequential dependence model for ad-hoc entity retrieval in the web of data

Zhiltsov N., Kotov A., Nikolaev F.

Kazan Federal University, 420008, Kremlevskaya 18, Kazan, Russia

Abstract

Previously proposed approaches to ad-hoc entity retrieval in the Web of Data (ERWD) used multi-fielded representation of entities and relied on standard unigram bag-of-words retrieval models. Although retrieval models incorporating term dependencies have been shown to be significantly more effective than the unigram bag-of-words ones for ad hoc document retrieval, it is not known whether accounting for term dependencies can improve retrieval from the Web of Data. In this work, we propose a novel retrieval model that incorporates term dependencies into structured document retrieval and apply it to the task of ERWD. In the proposed model, the document field weights and the relative importance of unigrams and bigrams are optimized with respect to the target retrieval metric using a learning-to-rank method. Experiments on a publicly available benchmark indicate significant improvement of the accuracy of retrieval results by the proposed model over state-of-the-art retrieval models for ERWD.

<http://dx.doi.org/10.1145/2766462.2767756>

Keywords

Entity retrieval, Knowledge graphs, Term dependence